

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 09/807,949C
Source: 1 Fw/b
Date Processed by STIC: 11/23/05

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 11/23/2005

PATENT APPLICATION: US/09/807,949C

TIME: 13:48:40

Input Set : A:\USSN 09-807,949-SEQ LISTING.txt

Output Set: N:\CRF4\11232005\I807949C.raw

3 <110> APPLICANT: Zavada, Jan
 4 Pastorekova, Silvia
 5 Pastorek, Jaromir
 7 <120> TITLE OF INVENTION: MN Gene and Protein
 9 <130> FILE REFERENCE: D-0021.5 PCT
 C--> 11 <140> CURRENT APPLICATION NUMBER: US/09/807,949C
 C--> 12 <141> CURRENT FILING DATE: 2001-08-09
 14 <150> PRIOR APPLICATION NUMBER: 09/177,776
 15 <151> PRIOR FILING DATE: 1998-10-23
 17 <150> PRIOR APPLICATION NUMBER: 09/178,115
 18 <151> PRIOR FILING DATE: 1998-10-23
 20 <160> NUMBER OF SEQ ID NOS: 143
 22 <170> SOFTWARE: PatentIn Ver. 2.1
 24 <210> SEQ ID NO: 1
 25 <211> LENGTH: 1522
 26 <212> TYPE: DNA
 27 <213> ORGANISM: HUMAN
 29 <220> FEATURE:
 30 <221> NAME/KEY: CDS
 31 <222> LOCATION: (13)..(1389)
 33 <220> FEATURE:
 34 <221> NAME/KEY: mat_peptide
 35 <222> LOCATION: (124)..(1389)
 37 <400> SEQUENCE: 1
 38 acagtcagcc gc atg gct ccc ctg tgc ccc agc ccc tgg ctc cct ctg ttg 51
 39 Met Ala Pro Leu Cys Pro Ser Pro Trp Leu Pro Leu Leu
 40 -35 -30 -25
 42 atc ccg gcc cct gct cca ggc ctc act gtg caa ctg ctg ctg tca ctg 99
 43 Ile Pro Ala Pro Ala Pro Gly Leu Thr Val Gln Leu Leu Leu Ser Leu
 44 -20 -15 -10
 46 ctg ctt ctg atg cct gtc cat ccc cag agg ttg ccc cgg atg cag gag 147
 47 Leu Leu Leu Met Pro Val His Pro Gln Arg Leu Pro Arg Met Gln Glu
 48 -5 -1 1 5
 50 gat tcc ccc ttg gga gga ggc tct tct ggg gaa gat gac cca ctg ggc 195
 51 Asp Ser Pro Leu Gly Gly Gly Ser Ser Gly Glu Asp Asp Pro Leu Gly
 52 10 15 20
 54 gag gag gat ctg ccc agt gaa gag gat tca ccc aga gag gag gat cca 243
 55 Glu Glu Asp Leu Pro Ser Glu Glu Asp Ser Pro Arg Glu Glu Asp Pro
 56 25 30 35 40
 58 ccc gga gag gag gat cta cct gga gag gag gat cta cct gga gag gag 291
 59 Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu
 60 45 50 55
 62 gat cta cct gaa gtt aag cct aaa tca gaa gaa gag ggc tcc ctg aag 339

P.6

RAW SEQUENCE LISTING

DATE: 11/23/2005

PATENT APPLICATION: US/09/807,949C

TIME: 13:48:40

Input Set : A:\USSN 09-807,949-SEQ LISTING.txt

Output Set: N:\CRF4\11232005\I807949C.raw

63	Asp	Leu	Pro	Glu	Val	Lys	Pro	Lys	Ser	Glu	Glu	Glu	Gly	Ser	Leu	Lys	
64				60					65					70			
66	tta	gag	gat	cta	cct	act	gtt	gag	gct	cct	gga	gat	cct	caa	gaa	ccc	387
67	Leu	Glu	Asp	Leu	Pro	Thr	Val	Glu	Ala	Pro	Gly	Asp	Pro	Gln	Glu	Pro	
68				75				80					85				
70	cag	aat	aat	gcc	cac	agg	gac	aaa	gaa	ggg	gat	gac	cag	agt	cat	tgg	435
71	Gln	Asn	Asn	Ala	His	Arg	Asp	Lys	Glu	Gly	Asp	Asp	Gln	Ser	His	Trp	
72				90				95					100				
74	cgc	tat	gga	ggc	gac	ccg	ccc	tgg	ccc	cgg	gtg	tcc	cca	gcc	tgc	gcg	483
75	Arg	Tyr	Gly	Gly	Asp	Pro	Pro	Trp	Pro	Arg	Val	Ser	Pro	Ala	Cys	Ala	
76	105					110				115						120	
78	ggc	cgc	ttc	cag	tcc	ccg	gtg	gat	atc	cgc	ccc	cag	ctc	gcc	gcc	ttc	531
79	Gly	Arg	Phe	Gln	Ser	Pro	Val	Asp	Ile	Arg	Pro	Gln	Leu	Ala	Ala	Phe	
80					125					130					135		
82	tgc	ccg	gcc	ctg	cgc	ccc	ctg	gaa	ctc	ctg	ggc	ttc	cag	ctc	ccg	ccg	579
83	Cys	Pro	Ala	Leu	Arg	Pro	Leu	Glu	Leu	Leu	Gly	Phe	Gln	Leu	Pro	Pro	
84				140					145				150				
86	ctc	cca	gaa	ctg	cgc	ctg	cgc	aac	aat	ggc	cac	agt	gtg	caa	ctg	acc	627
87	Leu	Pro	Glu	Leu	Arg	Leu	Arg	Asn	Asn	Gly	His	Ser	Val	Gln	Leu	Thr	
88				155				160					165				
90	ctg	cct	cct	ggg	cta	gag	atg	gct	ctg	ggg	ccc	ggg	cgg	gag	tac	cgg	675
91	Leu	Pro	Pro	Gly	Leu	Glu	Met	Ala	Leu	Gly	Pro	Gly	Arg	Glu	Tyr	Arg	
92		170					175				180						
94	gct	ctg	cag	ctg	cat	ctg	cac	tgg	ggg	gct	gca	ggg	cgt	ccg	ggc	tcg	723
95	Ala	Leu	Gln	Leu	His	Leu	His	Trp	Gly	Ala	Ala	Gly	Arg	Pro	Gly	Ser	
96	185					190				195					200		
98	gag	cac	act	gtg	gaa	ggc	cac	cgt	ttc	cct	gcc	gag	atc	cac	gtg	gtt	771
99	Glu	His	Thr	Val	Glu	Gly	His	Arg	Phe	Pro	Ala	Glu	Ile	His	Val	Val	
100					205					210				215			
102	cac	ctc	agc	acc	gcc	ttt	gcc	aga	gtt	gac	gag	gcc	ttg	ggg	cgc	ccg	819
103	His	Leu	Ser	Thr	Ala	Phe	Ala	Arg	Val	Asp	Glu	Ala	Leu	Gly	Arg	Pro	
104				220				225					230				
106	gga	ggc	ctg	gcc	gtg	ttg	gcc	gcc	ttt	ctg	gag	gag	ggc	ccg	gaa	gaa	867
107	Gly	Gly	Leu	Ala	Val	Leu	Ala	Ala	Phe	Leu	Glu	Glu	Gly	Pro	Glu	Glu	
108			235				240						245				
110	aac	agt	gcc	tat	gag	cag	ttg	ctg	tct	cgc	ttg	gaa	gaa	atc	gct	gag	915
111	Asn	Ser	Ala	Tyr	Glu	Gln	Leu	Leu	Ser	Arg	Leu	Glu	Glu	Ile	Ala	Glu	
112		250					255					260					
114	gaa	ggc	tca	gag	act	cag	gtc	cca	gga	ctg	gac	ata	tct	gca	ctc	ctg	963
115	Glu	Gly	Ser	Glu	Thr	Gln	Val	Pro	Gly	Leu	Asp	Ile	Ser	Ala	Leu	Leu	
116	265					270				275					280		
118	ccc	tct	gac	ttc	agc	cgc	tac	ttc	caa	tat	gag	ggg	tct	ctg	act	aca	1011
119	Pro	Ser	Asp	Phe	Ser	Arg	Tyr	Phe	Gln	Tyr	Glu	Gly	Ser	Leu	Thr	Thr	
120					285					290				295			
122	ccg	ccc	tgt	gcc	cag	ggg	gtc	atc	tgg	act	gtg	ttt	aac	cag	aca	gtg	1059
123	Pro	Pro	Cys	Ala	Gln	Gly	Val	Ile	Trp	Thr	Val	Phe	Asn	Gln	Thr	Val	
124				300				305					310				
126	atg	ctg	agt	gct	aag	cag	ctc	cac	acc	ctc	tct	gac	acc	ctg	tgg	gga	1107
127	Met	Leu	Ser	Ala	Lys	Gln	Leu	His	Thr	Leu	Ser	Asp	Thr	Leu	Trp	Gly	

RAW SEQUENCE LISTING

DATE: 11/23/2005

PATENT APPLICATION: US/09/807,949C

TIME: 13:48:40

Input Set : A:\USSN 09-807,949-SEQ LISTING.txt

Output Set: N:\CRF4\11232005\I807949C.raw

```

128          315          320          325
130 cct ggt gac tct cgg cta cag ctg aac ttc cga gcg acg cag cct ttg 1155
131 Pro Gly Asp Ser Arg Leu Gln Leu Asn Phe Arg Ala Thr Gln Pro Leu
132          330          335          340
134 aat ggg cga gtg att gag gcc tcc ttc cct gct gga gtg gac agc agt 1203
135 Asn Gly Arg Val Ile Glu Ala Ser Phe Pro Ala Gly Val Asp Ser Ser
136 345          350          355          360
138 cct cgg gct gct gag cca gtc cag ctg aat tcc tgc ctg gct gct ggt 1251
139 Pro Arg Ala Ala Glu Pro Val Gln Leu Asn Ser Cys Leu Ala Ala Gly
140          365          370          375
142 gac atc cta gcc ctg gtt ttt ggc ctc ctt ttt gct gtc acc agc gtc 1299
143 Asp Ile Leu Ala Leu Val Phe Gly Leu Leu Phe Ala Val Thr Ser Val
144          380          385          390
146 gcg ttc ctt gtg cag atg aga agg cag cac aga agg gga acc aaa ggg 1347
147 Ala Phe Leu Val Gln Met Arg Arg Gln His Arg Arg Gly Thr Lys Gly
148          395          400          405
150 ggt gtg agc tac cgc cca gca gag gta gcc gag act gga gcc 1389
151 Gly Val Ser Tyr Arg Pro Ala Glu Val Ala Glu Thr Gly Ala
152          410          415          420
154 tagaggctgg atcttgagaga atgtgagaag ccagccagag gcacctgagg gggagccggt 1449
156 aactgtcctg tcctgctcat tatgccactt ccttttaact gccaaagaat tttttaaaat 1509
158 aaatatttat aat 1522
161 <210> SEQ ID NO: 2
162 <211> LENGTH: 459
163 <212> TYPE: PRT
164 <213> ORGANISM: HUMAN
166 <400> SEQUENCE: 2
167 Met Ala Pro Leu Cys Pro Ser Pro Trp Leu Pro Leu Leu Ile Pro Ala
168          -35          -30          -25
170 Pro Ala Pro Gly Leu Thr Val Gln Leu Leu Leu Ser Leu Leu Leu Leu
171          -20          -15          -10
173 Met Pro Val His Pro Gln Arg Leu Pro Arg Met Gln Glu Asp Ser Pro
174          -5          -1 1          5          10
176 Leu Gly Gly Gly Ser Ser Gly Glu Asp Asp Pro Leu Gly Glu Glu Asp
177          15          20          25
179 Leu Pro Ser Glu Glu Asp Ser Pro Arg Glu Glu Asp Pro Pro Gly Glu
180          30          35          40
182 Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro
183          45          50          55
185 Glu Val Lys Pro Lys Ser Glu Glu Glu Gly Ser Leu Lys Leu Glu Asp
186          60          65          70          75
188 Leu Pro Thr Val Glu Ala Pro Gly Asp Pro Gln Glu Pro Gln Asn Asn
189          80          85          90
191 Ala His Arg Asp Lys Glu Gly Asp Asp Gln Ser His Trp Arg Tyr Gly
192          95          100          105
194 Gly Asp Pro Pro Trp Pro Arg Val Ser Pro Ala Cys Ala Gly Arg Phe
195          110          115          120
197 Gln Ser Pro Val Asp Ile Arg Pro Gln Leu Ala Ala Phe Cys Pro Ala
198          125          130          135

```

RAW SEQUENCE LISTING

DATE: 11/23/2005

PATENT APPLICATION: US/09/807,949C

TIME: 13:48:40

Input Set : A:\USSN 09-807,949-SEQ LISTING.txt

Output Set: N:\CRF4\11232005\I807949C.raw

```

200 Leu Arg Pro Leu Glu Leu Leu Gly Phe Gln Leu Pro Pro Leu Pro Glu
201 140          145          150          155
203 Leu Arg Leu Arg Asn Asn Gly His Ser Val Gln Leu Thr Leu Pro Pro
204          160          165          170
206 Gly Leu Glu Met Ala Leu Gly Pro Gly Arg Glu Tyr Arg Ala Leu Gln
207          175          180          185
209 Leu His Leu His Trp Gly Ala Ala Gly Arg Pro Gly Ser Glu His Thr
210          190          195          200
212 Val Glu Gly His Arg Phe Pro Ala Glu Ile His Val Val His Leu Ser
213          205          210          215
215 Thr Ala Phe Ala Arg Val Asp Glu Ala Leu Gly Arg Pro Gly Gly Leu
216 220          225          230          235
218 Ala Val Leu Ala Ala Phe Leu Glu Glu Gly Pro Glu Glu Asn Ser Ala
219          240          245          250
221 Tyr Glu Gln Leu Leu Ser Arg Leu Glu Glu Ile Ala Glu Glu Gly Ser
222          255          260          265
224 Glu Thr Gln Val Pro Gly Leu Asp Ile Ser Ala Leu Leu Pro Ser Asp
225          270          275          280
227 Phe Ser Arg Tyr Phe Gln Tyr Glu Gly Ser Leu Thr Thr Pro Pro Cys
228          285          290          295
230 Ala Gln Gly Val Ile Trp Thr Val Phe Asn Gln Thr Val Met Leu Ser
231 300          305          310          315
233 Ala Lys Gln Leu His Thr Leu Ser Asp Thr Leu Trp Gly Pro Gly Asp
234          320          325          330
236 Ser Arg Leu Gln Leu Asn Phe Arg Ala Thr Gln Pro Leu Asn Gly Arg
237          335          340          345
239 Val Ile Glu Ala Ser Phe Pro Ala Gly Val Asp Ser Ser Pro Arg Ala
240          350          355          360
242 Ala Glu Pro Val Gln Leu Asn Ser Cys Leu Ala Ala Gly Asp Ile Leu
243          365          370          375
245 Ala Leu Val Phe Gly Leu Leu Phe Ala Val Thr Ser Val Ala Phe Leu
246 380          385          390          395
248 Val Gln Met Arg Arg Gln His Arg Arg Gly Thr Lys Gly Gly Val Ser
249          400          405          410
251 Tyr Arg Pro Ala Glu Val Ala Glu Thr Gly Ala
252          415          420
256 <210> SEQ ID NO: 3
257 <211> LENGTH: 29
258 <212> TYPE: DNA
259 <213> ORGANISM: HUMAN
261 <400> SEQUENCE: 3
262 cgcccagtgg gtcattctcc ccagaagag
265 <210> SEQ ID NO: 4
266 <211> LENGTH: 19
267 <212> TYPE: DNA
268 <213> ORGANISM: HUMAN
270 <400> SEQUENCE: 4
271 ggaatcctcc tgcattccgg
274 <210> SEQ ID NO: 5

```

29

19

RAW SEQUENCE LISTING

DATE: 11/23/2005

PATENT APPLICATION: US/09/807,949C

TIME: 13:48:40

Input Set : A:\USSN 09-807,949-SEQ LISTING.txt

Output Set: N:\CRF4\11232005\I807949C.raw

275 <211> LENGTH: 10898
 276 <212> TYPE: DNA
 277 <213> ORGANISM: HUMAN
 279 <220> FEATURE:
 280 <221> NAME/KEY: gene
 281 <222> LOCATION: (1)..(10898)
 282 <223> OTHER INFORMATION: full-length MN genomic sequence
 284 <220> FEATURE:
 W--> 285 <221> NAME/KEY: unsure of base at position 1974
 286 <222> LOCATION: (1974)
 287 <223> OTHER INFORMATION: unsure of base at position 1974, which is in the 5' region
 flanking the transcription initiation site (3507) as determined by RNase protection assay.
 289 <400> SEQUENCE: 5
 290 ggatcctggt gactcgtgac cttacccccca accctgtgct ctctgaaaca tgagctgtgt 60
 291 ccactcaggg ttaaattgat taagggcggg gcaagatgtg ctttggttaa cagatgcttg 120
 292 aaggcagcat gctcgttaag agtcatcacc aatcccta atcaagtaat cagggacaca 180
 293 aacactgcgg aaggccgcag ggtcctctgc ctaggaaaac cagagacctt tgttcacttg 240
 294 tttatctgac cttccctcca ctattgtcca tgaccctgcc aaatccccct ctgtgagaaa 300
 295 caccacaaga ttatcaataa aaaaataaat ttaaaaaaaa aatacaaaaa aaaaaaaaaa 360
 296 aaaaaaaaaa gacttacgaa tagttattga taaatgaata gctattggta aagccaagta 420
 297 aatgatcata ttcaaaacca gacggccatc atcacagctc aagtctacct gatttgatct 480
 298 ctttatcatt gtcattcttt ggattcacta gattagtcac catcctcaaa attctcccc 540
 299 aagttcta atacgttccaa acatttaggg gttacatgaa gcttgaacct actaccttct 600
 300 ttgcttttga gccatgagtt gtaggaatga tgagtttaca ccttacatgc tggggattaa 660
 301 tttaaacttt acctctaagt cagttgggta gcctttggct tatttttgta gctaattttg 720
 302 tagttaatgg atgcactgtg aatcttgcta tgatagtttt cctccacact ttgccactag 780
 303 gggtaggtag gtactcagtt ttcagtaatt gcttacctaa gaccctaagc cctatttctc 840
 304 ttgtactggc ctttatctgt aatatgggca tatttaatac aatataat ttggagtttt 900
 305 tttgtttggt tgtttgtttg tttttttgag acggagtctt gcatctgtca tgcccaggct 960
 306 ggagtagcag tgggtgccatc tcggctcact gcaagctcca cctcccagat tcacgccatt 1020
 307 ttcctgcctc agcctcccga gtagctggga ctacaggcgc ccgccacat gcccggttaa 1080
 308 ttttttgtag ttttggtaga gacgggggtt caccgtgtta gccagaatgg tctcgatctc 1140
 309 ctgacttcgt gatccacccg cctcggcctc ccaaagttct gggattacag gtgtgagcca 1200
 310 ccgcacctgg ccaatttttt gagtctttta aagtaaaaat atgtcttgta agctggtaac 1260
 311 tatggtacat ttccttttat taatgtggtg ctgacgggtc tataggttct tttgagtttg 1320
 312 gcatgcatat gctacttttt gcagtccttt cattacattt ttctctcttc atttgaagag 1380
 313 catgttatat ctttttagctt cacttggtct aaaagggtct ctcattagcc taacacagtg 1440
 314 tcattggttg taccacttgg atcataagt gaaaaacagt caagaaattg cacagtaata 1500
 315 cttggttgta agagggatga ttcaggtgaa tctgacacta agaaactccc ctacctgagg 1560
 316 tctgagattc ctctgacatt gctgtatata ggcttttctt ttgacagcct gtgactgcgg 1620
 317 actatttttc ttaagcaaga tatgctaaag ttttggtgag ctttttccag agagaggtct 1680
 318 catatctgca tcaagtgaga acatataatg tctgcatgtt tccatatttc aggaatgttt 1740
 319 gcttggtgtt tatgctttta tatagacagg gaaacttgtt cctcagtgc ccaaaagagg 1800
 320 tgggaattgt tattggatat catcattggc ccacgctttc tgaccttggg aacaattaag 1860
 321 ggttcataat ctcaattctg tcagaattgg tacaagaaat agctgctatg tttcttgaca 1920
 W--> 322 ttccacttgg taggaaataa gaatgtgaaa ctcttcagtt ggtgtgtgtc cctngttttt 1980
 323 ttgcaatttc cttcttactg tgttaaaaaa aagtatgatc ttgctctgag aggtgaggca 2040
 324 ttcttaataca tgatctttta agatcaataa tataatcctt tcaaggatta tgtctttatt 2100
 325 ataataaaga taatttgtct ttaacagaat caataatata atcccttaaa ggattatatc 2160
 326 tttgctgggc gcagtggctc acacctgtaa tcccagcact ttgggtggcc aaggtggaag 2220

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/807,949C

DATE: 11/23/2005
TIME: 13:48:41

Input Set : A:\USSN 09-807,949-SEQ LISTING.txt
Output Set: N:\CRF4\11232005\I807949C.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:5; N Pos. 1974
Seq#:25; Xaa Pos. 3,4
Seq#:26; Xaa Pos. 3,4
Seq#:58; N Pos. 1968
Seq#:90; N Pos. 1968
Seq#:110; N Pos. 647

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:5; Line(s) 287
Seq#:9; Line(s) 521
Seq#:25; Line(s) 698,704
Seq#:26; Line(s) 719,724
Seq#:58; Line(s) 1362,1368,1369
Seq#:90; Line(s) 1803,1809,1810
Seq#:110; Line(s) 2093,2099

VERIFICATION SUMMARY

DATE: 11/23/2005

PATENT APPLICATION: US/09/807,949C

TIME: 13:48:41

Input Set : A:\USSN 09-807,949-SEQ LISTING.txt

Output Set: N:\CRF4\11232005\I807949C.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application Number
L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:285 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:5
L:322 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:1920
L:519 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:9
L:702 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:25
L:707 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25 after pos.:0
L:722 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:26
L:727 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26 after pos.:0
L:1366 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:58
L:1404 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:58 after pos.:1920
L:1801 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:90
L:1807 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:90
L:1845 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:90 after pos.:1920
L:2097 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:110
L:2113 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:110 after pos.:600